

Joseph R. Saveri (State Bar No. 130064)
 Steven N. Williams (State Bar No. 175489)
 Cadio Zirpoli (State Bar No. 179108)
 Christopher K.L. Young (State Bar No. 318371)
 Louis A. Kessler (State Bar No. 243703)
 Elissa A. Buchanan (State Bar No. 249996)
 Travis Manfredi (State Bar No. 281779)

JOSEPH SAVERI LAW FIRM, LLP

601 California Street, Suite 1000
 San Francisco, California 94108

Telephone: (415) 500-6800

Facsimile: (415) 395-9940

Email: jsaveri@saverilawfirm.com
 swilliams@saverilawfirm.com
 czirpoli@saverilawfirm.com
 cyoung@saverilawfirm.com
 lkessler@saverilawfirm.com
 eabuchanan@saverilawfirm.com
 tmanfredi@saverilawfirm.com

*Counsel for Individual and Representative
 Plaintiffs and the Proposed Class*

[Additional counsel on signature page]

**UNITED STATES DISTRICT COURT
 NORTHERN DISTRICT OF CALIFORNIA
 OAKLAND DIVISION**

J. DOE 1, et al.,
 Individual and Representative Plaintiffs,
 v.
 GITHUB, INC., et al.,
 Defendants.

Case Nos. 4:22-cv-06823-JST
 4:22-cv-07074-JST

**JOINT CASE MANAGEMENT
 STATEMENT**

Date: September 15, 2023
 Time: 1:30 p.m.
 Judge: Hon. Jon S. Tigar
 Ctrm: 6

Plaintiffs Does 1 through 5 and the proposed class (“Plaintiffs”); counsel for Defendants GitHub, Inc. (“GitHub”); Microsoft Corporation (“Microsoft”); and OpenAI, Inc.; OpenAI, LP; OpenAI OpCo, LLC; OpenAI GP, LLC; OpenAI Startup Fund GP I, LLC; OpenAI Startup Fund I, LP; and OpenAI Startup Fund Management, LLC (“OpenAI”) file this Joint Case Management Statement (“CMS”) pursuant to ECF No. 130.

I. STATUS OF DISCOVERY

A. Initial Disclosures

Pursuant to Federal Rule of Civil Procedure 26(a)(1)(A), Plaintiffs served their Initial Disclosures on June 30, 2023. Defendants¹ served their initial disclosures on July 5, 2023. The Parties will supplement Initial Disclosures, when necessary, as required by Rule 26(e).

B. Plaintiffs’ Discovery Requests

1. Discovery Propounded by Plaintiffs and Responses from Defendants.

Requests for Production of Documents (“RFPs”):

Plaintiffs’ Position: On July 12, 2023, Plaintiffs served First Sets of Requests for Production on GitHub (49 requests), Microsoft (60 requests), and OpenAI (57 requests). Defendants served responses and objections on August 18, 2023. As a general matter, Plaintiffs seek basic information regarding corporate structure; identification of those with management or other responsibility for the conduct alleged; the background and operation of the relevant technologies; the source and use of the training data and datasets; use of and compliance with applicable licenses; matches or near matches of plaintiffs’ data; submissions of information to regulators and governmental agencies; public statements regarding these topics; business and financial relationships between and among the parties and their principals; income, revenue and profits from sale or distribution of Copilot; and insurance policies and other materials ordinarily produced in cases of this type.² Broadly speaking, Defendants objected to all of Plaintiffs’ RFPs. Despite receiving an extension to respond to Plaintiffs’ RFPs, Defendants have yet to

¹ GitHub, Microsoft, and OpenAI are collectively referred to as “Defendants.”

² This general description is intended to generally describe the nature and scope of the document discovery sought by Plaintiffs to date. Plaintiffs and Defendants are meeting and conferring on the individual requests. Plaintiffs hope to bring this process to a close as soon as possible.

1 produce any documents. Only six RFPs for each Defendant could require disclosure of Plaintiffs' true
 2 names to anyone. For those RFPs, Plaintiffs granted Defendants an extra week to respond (a total of two
 3 weeks). Defendants served their response at the end of the one-week extension.

4 Consistent with the principles of cooperation underlying discovery as outlined by the Federal
 5 Rules, the Sedona Conference principles, and the Northern District of California's Guidelines for the
 6 Discovery of Electronically Stored Information, the Parties have begun the meet and confer process with
 7 respect to Defendants' document production. In the last week, Plaintiffs were able to have conducted
 8 initial face-to-face meet and confer sessions with GitHub, Microsoft, and OpenAI on all the discovery
 9 requests it has propounded thus far, conferring for approximately 2.5 hours for each Defendant.³
 10 Plaintiffs do not view this report as the occasion to litigate particular requests. Defendants apparently
 11 disagree and seek to raise issues with respect to particular requests, without providing the Court their
 12 objections and responses (which Plaintiffs assert do not comport with Rules 26 and 34) and without
 13 regard for the meet and confer calls which recently began.

14 Generally, Plaintiffs anticipate Defendants will produce responsive documents through (1)
 15 collection of responsive materials within their possession, custody, or control (*i.e.*, "go and get") and (2)
 16 implementing search terms for custodial ESI. Pursuant to an ESI protocol to be negotiated, certain
 17 responsive documents will be produced in native form. Plaintiffs anticipate certain issues will require
 18 court resolution. *See* Standing Order for All Civil Cases ("Civil Standing Order") § H.

19 **Defendants' Position:** Plaintiffs served discovery requests on July 12, 2023, following the
 20 July 11, 2023 Case Management Conference, during which Defendants raised the issue of Plaintiffs'
 21 designation of Plaintiffs' identities and code as HIGHLY CONFIDENTIAL—ATTORNEYS' EYES
 22 ONLY and the resulting impediment to Defendants' investigations. The Parties submitted a joint
 23 discovery letter on July 21, 2023 describing how Plaintiffs' confidentiality designations impair
 24 Defendants' ability to respond to Plaintiffs' discovery requests and litigate the case. (ECF No. 136.)
 25 The Court issued an order on July 26, 2023 rejecting Plaintiffs' proposal that Defendants' must disclose
 26 to Plaintiffs' counsel in advance the name of any person to whom Defendants wish to provide Plaintiffs'

27 _____
 28 ³ Plaintiffs met and conferred with respect to Microsoft and GitHub separately and individually for a
 total of approximately five hours.

names and source code. (ECF No. 138.) Given Defendants’ impaired abilities to investigate the case prior to the Court’s July 26, 2023 order, and the expansive breadth and number of requests, Defendants requested a 30-day extension to respond. Plaintiffs refused and granted only a one-week extension. Defendants served responses on August 18, 2023.

Many of Plaintiffs’ requests seek information entirely irrelevant to any claim or defense in this case, including, for example, requests for “ALL contracts, promissory notes, loans, obligations, estoppels, reservations of rights, investment agreements, shareholder agreements, shareholder rights agreements, purchase agreements, OR other agreements, including any drafts, between or among [DEFENDANT] and: a. ANY other DEFENDANT; b. ANY other officer or director of any DEFENDANT; or c. ANY entity substantially controlled by ANY officer or director of ANY other DEFENDANT, regardless of whether YOUR role is as a party, third-party beneficiary, OR other interested party” (*see, e.g.*, RFP No. 21 to OpenAI; RFP No. 23 to Microsoft and GitHub), regardless of whether such agreements have anything to do with Copilot or Codex; “ALL DOCUMENTS that pertain to the acquisition, purchase, or sale of an interest in OPENAI, INCLUDING due diligence materials exchanged or placed in data rooms in connection with any such transaction” and “IDENTIF[ICATION of] any third parties that participated in ANY acquisition, purchase, or sale of an interest in OPENAI, INCLUDING service providers, lawyers, investment banks, venture capitalists or consultants (*see, e.g.*, RFP Nos. 26-27 to OpenAI; Nos. 28-29 to Microsoft); “ALL DOCUMENTS YOU have provided to ANY United States regulatory body,” “United States Department of Justice or the Federal Trade Commission,” “the Securities and Exchange Commission,” “ANY regulatory body outside of the United States,” “ANY United States legislator or legislative body,” regardless of any relevance to Copilot or Codex (*see, e.g.*, RFP Nos. 28-33 to OpenAI; RFP Nos. 32-35 to Microsoft; RFP Nos. 27-31 to GitHub); and “DOCUMENTS CONCERNING the development of DALL-E-2” (RFP No. 48 to Microsoft).

Notwithstanding the irrelevance and overbreadth of most of Plaintiffs’ document requests, Defendants still agreed to produce documents in response to many requests, agreed to meet and confer on others, and stated that no responsive documents exist for others. For instance, out of 49 requests propounded on GitHub, it agreed to produce documents in response to 23 requests and stated that it does

1 not have responsive documents (or has no reason to believe it has responsive documents) in response to
2 9 requests. Similarly, out of 59 requests propounded on Microsoft, it agreed to produce documents in
3 response to 16 requests and stated that it does not have responsive documents (or has no reason to
4 believe it has responsive documents or that responsive documents are in the possession of another
5 defendant) in response to 23 requests. Out of 57 requests propounded on OpenAI, OpenAI agreed to
6 produce documents or further meet and confer regarding 20 requests. Plaintiffs propounded largely the
7 same requests on all Defendants regardless of whether the request made sense for a particular
8 Defendant. Consequently, each Defendant responded in responses or during meet and confers that it
9 does not have responsive documents to a number of requests.

10 On August 22, 2023, Plaintiffs requested a meet and confer with Microsoft and GitHub's
11 counsel. Microsoft and GitHub's counsel requested that Plaintiffs send some writing prior to the meet
12 and confer outlining any concerns with the responses so that Defendants could prepare for the meeting.
13 Plaintiffs refused. Plaintiffs' counsel then conducted a 2-hour meeting about GitHub's RFP responses .
14 OpenAI counsel attended the meet and confer, after which Plaintiffs asked for a separate meet and
15 confer with OpenAI. Again, without any writing outlining Plaintiffs' concerns regarding OpenAI's
16 responses, the Parties met and conferred on September 5, 2023 for 2.5 hours. On September 7, 2023, the
17 Parties met and conferred for another 2.5 hours regarding Defendants' responses.

18 Defendants are in the process of determining the location of responsive non-custodial documents
19 and are collecting those documents for production. For instance, counsel for Defendants have
20 interviewed dozens of individuals to investigate potentially responsive material and have identified
21 multiple sources of noncustodial information. Defendants are hampered in collecting custodial
22 information because the Plaintiffs did not propose a written ESI Protocol until the evening of September
23 5, were not prepared to discuss that proposal in any of the prior lengthy meetings, and the parties have
24 not yet discussed how many and which custodians and keywords reflect a proportional scope of
25 production in this case.

26 **Interrogatories:**

27 **Plaintiffs' Position:** On July 19, 2023, Plaintiffs propounded 14 interrogatories each on GitHub,
28 Microsoft, and OpenAI. The interrogatories, among other things, ask Defendants to identify persons

1 responsible for the subject matter or events relevant to Plaintiffs' claims, description or identification of
 2 training data, and identification and location of current and former employees who are likely to possess
 3 relevant information. None of Plaintiffs interrogatories could require disclosure of Plaintiffs' true names
 4 to anyone. Following an extension, on August 25, 2023, Defendants provided limited responses to some
 5 interrogatories, indicating their investigation was not complete, and objected to others. Initial meet and
 6 confer sessions have been conducted for all outstanding interrogatories propounded by Plaintiffs.
 7 Plaintiffs hope to resolve the open issues without court resolution.

8 **Defendants' Position:** Defendants sought a 30-day extension to respond to Plaintiffs' July 19,
 9 2023 interrogatories given the breadth of the information Plaintiffs sought and because of Defendants'
 10 impaired ability to investigate as a result of Plaintiffs' confidentiality designations—an issue the Court
 11 resolved on July 26, 2023. Plaintiffs refused and agreed to only a one-week extension. Defendants
 12 responded to Plaintiffs' first set of interrogatories on August 25, 2023. The Parties are continuing to
 13 meet and confer regarding their responses.

14 Like Plaintiffs' RFPs, many of Plaintiffs' interrogatories seek information irrelevant to any claim
 15 or defense. For example, Plaintiffs seek identification of all of Defendants' past and present directors
 16 and officers (Interrogatory No. 2) and all individuals who have ever possessed stock or ownership
 17 interest in any Defendant greater than five percent (Interrogatory No. 3). These requests are irrelevant to
 18 all Defendants, but are also especially burdensome for a company like Microsoft, which has been in
 19 existence for nearly 50 years and is publicly traded. Still, Microsoft and GitHub substantively
 20 responded to 11 of the 14 interrogatories. Likewise, OpenAI responded to 7 of the 14 interrogatories,
 21 while several other interrogatories sought information not in OpenAI's possession.

22 **2. Discovery Discussed with Defendants But Not Propounded:**

23 **Plaintiffs' Position:** Defendant GitHub, citing generalized concerns with cost and burden, have
 24 indicated certain ESI—namely so-called “telemetry data”—will soon be destroyed. Plaintiffs discussed
 25 this issue with GitHub when it first informed Plaintiffs of its plans for deletion and deferred a more
 26 detailed discussion in anticipation of specific request for production. GitHub has confirmed that
 27 telemetry data is responsive to one or more of Plaintiffs' requests for production. Plaintiffs have
 28 requested GitHub articulate with specificity the burden associated with preserving that ESI. GitHub has

1 yet to provide an estimate of the cost with any specificity nor an estimate of the storage capacity taken
 2 up by each of the first six months of data slated for deletion. Similar highly relevant ESI was slated for
 3 destruction by OpenAI. Plaintiffs are gratified to have just been been informed that OpenAI has
 4 suspended any scheduled destruction and are hopeful these issues can be resolved collaboratively. As
 5 with GitHub, OpenAI has not yet quantified the burden resulting from this data. Plaintiffs have
 6 requested Defendants delay any destruction. Plaintiffs note that the case cited by Defendants in support
 7 of their vague burden claim is plainly distinguishable. In *Rodriguez v. Google LLC*, Google was
 8 collecting “**over 27 petabytes” per day** compared to **0.5 petabytes per year** by GitHub here. 2021 WL
 9 8085492, at *1.

10 Plaintiffs have informed Defendants they will likely propound further interrogatories generally
 11 and eventually will seek to take depositions of present and former employees. This will likely require
 12 expansion of the number of depositions set forth in Rule 30. Plaintiffs also anticipate the depositions of
 13 company designee(s) pursuant to Rule 30(b)(6). In addition, Plaintiffs have informed Defendants they
 14 intend to seek discovery from nonparties using subpoenas pursuant to Federal Rule of Civil Procedure
 15 45. Plaintiffs also anticipate expert disclosures and expert depositions pursuant to Rule 26.

16 **Defendants’ Position:** As stated in the Parties’ July 5, 2023 Case Management Statement (ECF
 17 No. 116), Defendants alerted Plaintiffs to certain upcoming data deletion and the Parties met and
 18 conferred regarding whether Defendants will preserve such data. Defendant GitHub first alerted
 19 Plaintiffs on June 1, 2023 that a petabyte-sized collection of two years’ worth of certain types of
 20 Copilot user data would begin rolling off of servers in July and September, 2023, respectively. GitHub
 21 also explained that it would maintain two years’ worth of such data on a rolling basis, that it was not
 22 technically possible to export the expiring data collection, that the rate of growth in the size of the data
 23 collection is significant, that it was extremely expensive to maintain the data indefinitely, and that there
 24 was no reason it was aware of to believe that such data was materially different in any relevant way
 25 from the earlier data that it would no longer maintain. *See Rodriguez v. Google LLC*, No. 20-CV-
 26 04688-RS (AGT), 2021 WL 8085492, at *1-2 (N.D. Cal. Dec. 1, 2021) (sampling of extremely large
 27 rolling dataset sufficient). GitHub again identified the issue in the Parties’ July 5, 2023 Case
 28 Management Statement (ECF No. 116). GitHub agreed to temporarily maintain the first sixty days of

1 data for an additional sixty days to allow Plaintiffs to make a motion requiring preservation of the data
 2 if they wish to do so. The parties are meeting and conferring on the issue. Similarly, OpenAI alerted
 3 Plaintiffs on June 1, 2023 that per public announcements and press coverage, OpenAI had updated its
 4 retention of Codex API data. After the Parties met and conferred on June 16, 2023, at Plaintiffs’
 5 request, OpenAI suspended deletion of prompts and generations for its Codex API. There is no live
 6 issue requiring Court intervention, as OpenAI has acceded to Plaintiffs’ request to suspend its retention
 7 policy for the relevant data.

8 Plaintiffs indicate that they anticipate seeking discovery from nonparties using subpoenas, but
 9 have not identified any of these individuals nor provided a justification for expanding the number of
 10 depositions.

11 **C. Defendants’ Discovery Requests**

12 **1. Discovery Propounded by Defendants and Responses from Plaintiffs** 13 **Requests for Production of Documents (“RFPs”):**

14 **Plaintiffs’ Position:** On July 12, 2023, OpenAI served a separate set of RFPs on each of the five
 15 plaintiffs. On August 11, 2023, all Plaintiffs served objections and responses to OpenAI’s first sets of
 16 RFPs and produced responsive documents (131 pages). Plaintiffs’ production is complete. On July 27,
 17 2023, GitHub and Microsoft jointly served separate sets of RFPs on each of the five Plaintiffs, broadly
 18 seeking material related to threats that Plaintiffs’ received (the “Joint RFPs”). On August 28, 2023,
 19 Plaintiffs served objections and responses to the Joint RFPs. Specifically, Plaintiffs objected and refused
 20 to produce in response to Requests 2 through 5 on the grounds that the RFPs seek information protected
 21 by attorney client privilege; are duplicative of OpenAI’s RFPs and interrogatories; and have been the
 22 subject of law and motion practice and resolved by the Court (*See* ECF Nos. 71, 77, 77-1, 77-2, 77-3,
 23 77-4, 77-5, 77-6, and 77-7). Plaintiffs are prepared to meet and confer on their responses should
 24 Defendants ask them to.

25 To the extent Defendants accuse Plaintiffs of deleting relevant ESI, the relevant material is either
 26 preserved or in Defendants’ care, custody, and control. In any event, Plaintiffs are willing to meet and
 27 confer regarding Defendants’ concerns, but no such request has been made.

Defendants' Position: In response to 24 RFPs, Plaintiffs have produced 43 documents. These documents consist of (1) four unreadable JSON files that appear to be files from an unknown messaging platform; (2) one blank HTML file; (3) 14 emails without any metadata (that do not appear to be relevant to the litigation), (4) 15 PNG files that appear to be pictures of, presumably, Slack, twitter, and text message communications from three of the Plaintiffs discussing the case; and (5) several excerpted comment sections from a website called Hacker News. Plaintiffs assert their production is complete, but Plaintiffs have not produced many of the documents they agreed to produce in response to Defendants' RFPs. For instance, Plaintiffs agreed to, but have not, produced any documents relating to Plaintiffs' creation or ownership of their allegedly copyrighted works, copies of any of their allegedly copyrighted works, documents related to their decision to enter into any licenses they allege apply to their allegedly copyrighted works, any copyright management information associated with any of their alleged copyrighted works, any agreements related to the alleged copyrighted works, any documents relating to any injury or harm Plaintiffs claim to have suffered, and any documents relating to Plaintiffs' efforts to enforce their allegedly copyrighted works.

GitHub is also concerned because once it was able to identify the Plaintiffs to employees who could assist in investigating their accounts, its initial investigation revealed that one or more of the Plaintiffs destroyed public code repositories either shortly prior to or months after the filing of this lawsuit. Plaintiffs' refusal to identify themselves in the pleading, and the refusal thereafter to permit disclosure of their names to non-lawyer personnel, prevented GitHub from freezing their data under a litigation hold and thus directly interfered with GitHub's ability to prevent this destruction from happening. Further, in response to requests seeking all documents supporting Plaintiffs' contentions that the disclosure of their identities would threaten their safety, Plaintiffs refused to produce on the grounds that the matter has been resolved by the Court. The issue of Names Information is going to keep coming before the Court at multiple stages of the proceeding. *See Part II. 1., infra.*

Interrogatories:

Plaintiffs' Position: On July 12, 2023, OpenAI served a separate set of interrogatories on each of the five Plaintiffs. On August 11, 2023, Plaintiffs served objections and complete responses to all of

OpenAI's interrogatories. Plaintiffs are prepared to meet and confer on their responses should OpenAI ask them to.

Defendants' Position:

OpenAI intends to meet and confer with Plaintiffs regarding deficiencies in their responses.

2. Discovery Discussed with Plaintiffs But Not Propounded

Defendants have asserted a need to send subpoenas to nonparties. But, at present, Plaintiffs are not aware of any particular nonparties Defendants intend to serve with discovery requests.

II. OTHER ISSUES

1. Issues Requiring the Court's Attention

Discussion of First Amended Complaint at Upcoming Hearing:

Plaintiffs' Position: On August 22, 2023, Defendants GitHub and Microsoft informed Plaintiffs they intend to disclose certain details of Plaintiffs' code in open court at the upcoming Motion to Dismiss hearing initially set for September 14, 2023 (ECF Nos. 108, 110), but recently continued to November 9, 2023 (ECF No. 148). Disclosure of Plaintiffs' code in connection with Plaintiffs' identities as Doe plaintiffs in this action could potentially expose their identities. This information was ordered to be filed under seal in this case. Plaintiffs responded in writing to Defendants request on August 31, 2023 and are prepared to meet and confer. If agreement cannot be reached, the Parties will be prepared to discuss GitHub/Microsoft's request at the Case Management Conference.

Defendants' Position: Plaintiffs seriously mischaracterize GitHub and Microsoft's position. GitHub and Microsoft made clear that they do *not* intend to disclose the detailed contents of the source code, but rather to refer at a very high level to the overall functionality of that code as alleged in the First Amended Complaint. Specifically, in a letter dated August 22, 2023, GitHub and Microsoft proposed that in their view it was consistent with the Protective Order and would not reveal Names Information if they used the following phrases at the hearing on the Motion to Dismiss the Second Corrected First Amended Complaint: (1) placing pieces in their initial starting positions on a chess board, (2) listing the 20 amino acids or basic chemical building blocks, and (3) containing tests for code that validates network hostnames. Defendants wish to use these words both in order to respond to Plaintiffs' contentions about the plausibility of output copying and to illustrate certain important points regarding

the interpretation of 17 U.S.C. §1202. Such references are consistent with the Protective Order and the Court’s prior orders, since this abstract description would not tend to identify any Plaintiff.

Plaintiffs insisted that even such abstract descriptions would require sealing the courtroom and the only option was to do so or to limit argument to refer solely to paragraph numbers of the pleading. GitHub and Microsoft do not believe it appropriate to seal the courtroom, nor should they be restricted in this fashion. Sealing must be evaluated at “each stage of the proceedings.” *Berkley Ins. Co. v. OakCare Med. Grp.*, No. 22-CV-05181-HSG, 2023 WL 2804380, at *1 (N.D. Cal. Apr. 4, 2023) (quoting *Does I through XXIII v. Advanced Textile*, 214 F.3d 1058, 1068 (9th Cir. 2000)). Referencing the code in this high-level manner at the hearing fairly balances Plaintiffs’ requested anonymity against GitHub and Microsoft’s ability to fully advocate their positions and the public’s First Amendment right to open judicial proceedings.

Retention of Inaccessible ESI:

Plaintiffs’ Position: As discussed above, GitHub and OpenAI have indicated certain relevant ESI is too costly and burdensome to maintain and preserve, and thus will be deleted and/or is already being deleted. Specifically, GitHub has indicated telemetry data from two years and two months ago—including “the prompts input by the user and the code output suggestions that result therefrom”—are scheduled to be destroyed beginning September 4, 2023 on a rolling basis. OpenAI states in its June 7, 2023 letter “effective March 1, 2023, any data sent through the API will be retained for abuse and misuse monitoring purposes for a maximum of 30 days, after which it will be deleted (unless otherwise required by law). That API data includes user prompts and completions” (internal quotation marks omitted). It is unclear whether any of OpenAI’s telemetry data has been retained from before 30-days ago. Notably, this deletion policy went into effect over six weeks after this case was filed and OpenAI only notified Plaintiffs it was deleting crucially important data over three months later.

No Party disputes this ESI is relevant to issues in the case—perhaps even essential to determine whether Plaintiffs’ allegations are accurate. Plaintiffs have requested samplings of the ESI and that Defendants articulate with specificity the burden associated with maintaining that ESI—specifically, the cost for each of the first six months and the amount of data collected for the same period. Thus far,

Defendants have provided no samples or further details. If compromise cannot be reached, Court intervention will likely be required in the near future to avoid losing more of this data forever.

Defendants' Position: Defendant GitHub first alerted Plaintiffs on June 1, 2023 that a petabyte-sized collection of two years' worth of certain types of Copilot user data would begin rolling off of servers in July and September, 2023, respectively. GitHub also explained that it would maintain two years' worth of such data on a rolling basis, that it was not technically possible to export the expiring data collection, that the rate of growth in the size of the data collection is significant, that it was extremely expensive to maintain the data indefinitely, and that there was no reason it was aware of to believe that such data was materially different in any relevant way from the earlier data that it would no longer maintain. *See Rodriguez v. Google LLC*, No. 20-CV-04688-RS (AGT), 2021 WL 8085492, at *1-2 (N.D. Cal. Dec. 1, 2021) (sampling of extremely large rolling dataset sufficient). GitHub again identified the issue in the Parties' July 5, 2023 Case Management Statement (ECF No. 116). Still, GitHub agreed to temporarily maintain the first sixty days of data for an additional sixty days to allow Plaintiffs to make a motion requiring preservation of the data if they wish to do so. The parties are meeting and conferring on the issue.

As discussed above, OpenAI alerted Plaintiffs on June 1, 2023 that per public announcements and press coverage, OpenAI had updated its retention of Codex API data. After the Parties met and conferred on June 16, 2023, at Plaintiffs' request, OpenAI suspended deletion of prompts and generations for its Codex API. There is no live issue requiring Court intervention, as OpenAI has acceded to Plaintiffs' request to suspend its retention policy for the relevant data.

As discussed during the Parties' meet and confers, Defendants are open to further discussing how to query such data with Plaintiffs.

Stipulated Orders: Plaintiffs have recently prepared and circulated drafts of the following case management discovery agreements with Defendants: (1) an ESI Protocol, circulated on the evening of September 5, 2023; (2) a Rule 502(d) Stipulation, circulated on August 25, 2023; and (3) an expert discovery stipulation, circulated on August 28, 2023.

Discovery Limits and Deposition Locations:

Plaintiffs' Position: As discussed in the Parties' Rule 26(f) Report (ECF No. 116), this case will likely merit expansion of the discovery limits set forth in the Federal Rules of Civil Procedure.

Defendants' Position: Plaintiffs have not justified any need for an expansion of the discovery limits.

2. Issues Bearing on the Progress of the Case**Discovery Schedule.**

Plaintiffs' Position: In order to advance the litigation in an orderly and efficient manner, Plaintiffs would request that interim deadlines be set for completion of the meet and confers regarding the outstanding document requests and interrogatories and an initial deadline for the commencement of production of responsive documents by Defendants. Plaintiffs request the production begin within 21 days. Plaintiffs would also request a deadline for the completion of initial negotiations of a procedure and process for custodial ESI, including search terms. Plaintiffs request the negotiations be ordered completed within 30 days.

Defendants' Position: Plaintiffs provided a draft ESI order to Defendants for the first time on the evening of September 5, 2023 and Defendants are considering it. Given Defendants' impaired ability to investigate this case prior to the Court's July 26, 2023 discovery order, Defendants' investigations are ongoing. As stated in Defendants' responses, Defendants will produce documents after an ESI order has been entered.

OpenAI Global, LLC:

Plaintiffs' Position: Plaintiffs plan to add OpenAI Global, LLC as a defendant at the soonest appropriate time. OpenAI Global, LLC is an entity Plaintiffs recently discovered. It is not identified in any corporate disclosure forms filed in this case nor in OpenAI's Initial Disclosures. However, it is the parent company of Defendant OpenAI OpCo, LLC. Microsoft owns at least 10% of OpenAI OpCo, LLC and it is a key part of OpenAI's corporate structure. Plaintiffs have informed OpenAI's counsel they intend to add this entity as a defendant and asked for an explanation as to why the entity was not identified in any of OpenAI's corporate disclosure forms or Initial Disclosures.

OpenAI Global, LLC is the only for-profit entity in the OpenAI organization (other than, perhaps, OpenAI OpCo, LLC). It was incorporated December 26, 2022, over six weeks after this case began. It apparently takes the place in OpenAI's structure previously occupied by Defendant OpenAI, LP, which became Defendant OpenAI OpCo, LLC sometime after the initial complaints in this case were filed. Plaintiffs are still awaiting a response to their August 23, 2023 email message requesting further information from OpenAI's counsel.

Defendants' Position: On March 21, 2023, OpenAI counsel emailed Plaintiffs' counsel informing them that OpenAI, L.P. had recently changed its name to OpenAI OpCo, LLC. OpenAI sought permission from Plaintiffs to file a stipulation regarding the name change to the Court, the draft of which included a certificate of conversion showing the change. Plaintiffs did not respond. OpenAI followed up on April 4, 2023. Plaintiffs requested additional information regarding corporate documents. OpenAI reattached the certificate of conversion. On April 18, 2023, OpenAI provided Plaintiffs with additional documents on file with the Delaware Secretary of State pertaining to OpenAI OpCo, LLC., including the certificate of formation. Plaintiffs did not respond. OpenAI followed up with Plaintiffs again on April 25, 2023. Plaintiffs did not respond. Instead of responding to OpenAI, Plaintiffs filed a First Amended Complaint on June 8, 2023 adding OpenAI OpCo, LLC to the complaint, but not removing OpenAI, L.P., despite the fact that OpenAI, L.P. no longer exists, having converted to become OpenAI OpCo, LLC.

Plaintiffs' approach appears to be to name and add all OpenAI entities it is aware of, regardless of whether the particular OpenAI entity is in any way involved in Codex or Copilot, or in possession of or has control over relevant documents. OpenAI explained to Plaintiffs during the Parties' September 5, 2023 meet and confer that the OpenAI Fund entities do not have any involvement in the claims and defenses at issue in this case. OpenAI is open to further meeting and conferring with Plaintiffs regarding which OpenAI entities are actually relevant to the case and to narrow the named OpenAI entities.

1 Dated: September 8, 2023

2 By: /s/ Joseph R. Saveri
3 Joseph R. Saveri

4 Joseph R. Saveri (State Bar No. 130064)
5 Steven N. Williams (State Bar No. 175489)
6 Cadio Zirpoli (State Bar No. 179108)
7 Christopher K.L. Young (State Bar No. 318371)
8 Louis A. Kessler (State Bar No. 243703)
9 Elissa A. Buchanan (State Bar No. 249996)
10 Travis Manfredi (State Bar No. 281779)
11 **JOSEPH SAVERI LAW FIRM, LLP**
12 601 California Street, Suite 1000
13 San Francisco, California 94108
14 Telephone: (415) 500-6800
15 Facsimile: (415) 395-9940
16 Email: jsaveri@saverilawfirm.com
17 swilliams@saverilawfirm.com
18 czirpoli@saverilawfirm.com
19 cyoung@saverilawfirm.com
20 lkessler@saverilawfirm.com
21 eabuchanan@saverilawfirm.com
22 tmanfredi@saverilawfirm.com

23 Matthew Butterick (State Bar No. 250953)
24 1920 Hillhurst Avenue, #406
25 Los Angeles, CA 90027
26 Telephone: (323) 968-2632
27 Facsimile: (415) 395-9940
28 Email: mb@buttericklaw.com

Counsel for Plaintiffs and the Proposed Class

1 Dated: September 8, 2023

By: /s/ Annette L. Hurst
Annette L. Hurst

3 Annette L. Hurst (State Bar No. 148738)
4 **ORRICK, HERRINGTON & SUTCLIFFE LLP**
5 The Orrick Building
6 405 Howard Street
7 San Francisco, CA 94105
8 Telephone: 415-773-5700
9 Facsimile: 415-773-5759
10 Email: ahurst@orrick.com

11 *Counsel for GitHub, Inc. and Microsoft Corporation*

12 Dated: September 8, 2023

By: /s/ Joseph C. Gratz
Joseph C. Gratz

13 Joseph C. Gratz (State Bar No. 240676)
14 **MORRISON & FOERSTER LLP**
15 425 Market Street
16 San Francisco, CA 94105
17 Telephone: 415-268-7000
18 Facsimile: 415-268-7522
19 Email: jgratz@mofo.com

20 *Counsel for OpenAI, Inc., OpenAI, L.P., OpenAI GP,*
21 *LLC, OpenAI OpCo, L.L.C., OpenAI Startup Fund GP*
22 *I, L.L.C, OpenAI Startup Fund I, L.P., and OpenAI*
23 *Startup Fund Management LLC*

L.R. 5-1 SIGNATURE ATTESTATION

As the ECF user whose user ID and password are utilized in the filing of this document, I attest under penalty of perjury that concurrence in the filing of the document has been obtained from each of the other signatories.

Dated: September 8, 2023

/s/ Joseph R. Saveri

Joseph R. Saveri